University of California, Berkeley

Physics 110B Fall 2003 (*Strovink*)

Excerpts from **Special Relativity Notes** (Physics 110B, S'97)

I. Lorentz Transformation

Notation. Our notation will be:

$\vec{A}$ is a 3-vector (A^1, A^2, A^3), with component $A^m, 1 \le m \le 3$.
$|\vec{A}|^2$ is its magnitude squared: $|\vec{A}|^2 = \vec{A} \cdot \vec{A} = A^1A^1 + A^2A^2 + A^3A^3$.

A is a 4-vector (A^0, A^1, A^2, A^3), with component A^μ, $0 \le \mu \le 3$.
A^2 is its magnitude squared: $A^2 = A \cdot A = A^0A^0 - A^1A^1 - A^2A^2 - A^3A^3$.

$\mathcal{A}$ is a 4-tensor, with 16 components A^μ_ν, $0 \le \mu, \nu \le 3$. In matrix notation, the top index is the row index and the bottom index is the column index. The product of $\mathcal{A}$ with a 4-vector B is denoted by $\mathcal{A} \cdot B$. To take that product, write B as a column vector and follow the usual rules of matrix algebra.

For those of you who have never seen components of 4-vectors written as superscripts before, please ignore this paragraph. For the rest of you, some comments on my notation might be useful. I am using a timelike-**positive** metric, with contravariant 4-vectors. Tensors have mixed indices (first index contravariant, second index covariant). The usual metric tensor is used implicitly when taking the product of two contravariant 4-vectors, but is not needed when taking the product of a mixed-index tensor with a 4-vector. In neither case is the metric tensor explicitly displayed. Thus, my notation is consistent with that used in standard texts on relativistic quantum mechanics, *e.g.* Bjorken & Drell, but (for pedagogic simplicity) I put the least possible emphasis on transforming between covariant and contravariant indices. Griffiths' notation is different in that (*i*) Griffiths uses a timelike-**negative** metric, changing the overall sign of the product of two 4-vectors; (*ii*) he puts more emphasis on transforming between contravariant and covariant indices. My notation is consistent with that of Jackson.

Quantities will be denoted as unprimed in the laboratory frame, and primed in a frame moving in the $\hat{x}$ or $\vec{1}$ direction with relative velocity $\beta_0 c$, where c is the speed of light. We also use $\gamma_0 \equiv (1 - \beta_0^2)^{-1/2}$.

Galilei Transformation. In this notation, the nonrelativistic transformation that you have been using since the eighth grade is:

$$\begin{aligned} ct &= ct' \ \text{ or } \ x^0 = x'^0 \\ x^1 &= x'^1 + \beta_0 ct' \\ x^2 &= x'^2 \\ x^3 &= x'^3 \ . \end{aligned}$$

By defining $x^0 = ct$, I choose to consider $x = (ct, x^1, x^2, x^3) = (ct, \vec{x})$ to be a 4-vector. Then the Galilei transformation becomes

$$x = \Lambda_g \cdot x' \ .$$

Exercise 1a. Evaluate Λ_g, the matrix for the Galilei transformation.

The Galilei transformation is obviously unsuited to the equations of electromagnetism. We have seen that the general solution to the homogeneous wave equation

$$\left(\nabla^2 - \frac{\partial^2}{c^2 \partial t^2}\right)\psi = 0 \ ,$$

where ψ is a component of $\vec{E}$ or $\vec{B}$, is a wave with phase velocity c, irrespective of the coordinate system used. The Galilei transformation instead implies the linear addition of velocities, leading to different wave velocities c in different coordinate systems. Still, we seek a transformation that reduces to the Galilei transformation in the nonrelativistic limit.

Derivation of the Lorentz Transformation. We take the two coordinate systems to be coincident at $t = t' = 0$. If a spherical light pulse is emitted from the origin(s) at that time, we have for the pulse in each system

$$(x^0)^2 = (x^1)^2 + (x^2)^2 + (x^3)^2$$
$$(x'^0)^2 = (x'^1)^2 + (x'^2)^2 + (x'^3)^2 \ .$$

We shall specialize to possible *linear* transformations. (As it happens, a rigorous proof that the Lorentz transformation must be linear is rather involved.)

Consider a rod with one end at the origin. If the rod is transverse to the direction of motion, we expect both ends to be coincident at the same time. Thus, for example, x^3 must be proportional to x'^3. Reciprocity requires the constant of proportionality to be unity.

Thus we have reduced the transformation to

$$\begin{aligned} x^0 &= g(x'^0 + bx'^1) \\ x^1 &= \gamma(x'^1 + \beta x'^0) \\ x^2 &= x^2 \\ x^3 &= x^3 \ , \end{aligned}$$

where g, b, γ, and β are as yet unspecified constants.

Exercise 1b. Complete the above derivation, by showing that $g = \gamma$, $b = \beta$, and $\gamma = (1-\beta^2)^{-1/2}$. To conform to the Galilean transformation in the nonrelativistic limit, we identify $\beta = \beta_0$, the relative velocity $\div c$.

Exercise 1c. Using the result of Exercise 2 and the notation of Exercise 1, show that the Lorentz transformation matrix defined by $x = \Lambda \cdot x'$ is given by

$$\Lambda = \begin{pmatrix} \gamma_0 & \gamma_0\beta_0 & 0 & 0 \\ \gamma_0\beta_0 & \gamma_0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \ .$$

Also show that Λ reduces to the Galilean transformation Λ_g for $\beta_0 \ll 1$.

Exercise 2a. If x is a 4-vector, that is if x transforms according to the Lorentz transformation, show by direct calculation that $x \cdot x = x' \cdot x'$. (Thus the dot product of these two 4-vectors is a Lorentz invariant. This is true for any two 4-vectors.)

Exercise 2b. Show by direct calculation that $\Lambda(-\beta_0) \cdot \Lambda(\beta_0) = \mathcal{I}$, where $\mathcal{I}$ is the unit matrix, and the dot signifies matrix multiplication. (Evidently, $\Lambda(-\beta_0)$ is the *inverse* Lorentz transformation: $x' = \Lambda(-\beta_0) \cdot x$. So if you transform from the primed frame to the lab frame, and then back to the primed frame, you recover the original 4-vector.)

Spacetime Intervals. The Lorentz transformation is an orthogonal transformation in the 4-dimensional space called *spacetime.* Let x_a and x_b denote the coordinates in spacetime of events a and b. The 4-vector interval $\Delta x = x_a - x_b$ between these two events has a Lorentz-invariant magnitude squared equal to

$$\begin{aligned} \Delta x^2 = (x_a^0 - x_b^0)^2 - \\ - (x_a^1 - x_b^1)^2 - (x_a^2 - x_b^2)^2 - (x_a^3 - x_b^3)^2 \ . \end{aligned}$$

If Δx^2 is positive, that is if the first (time) term exceeds the sum of the last 3 (space) terms, the interval is said to be *timelike.* If the opposite, the interval is *spacelike.* If $\Delta x^2 = 0$, the interval is called *lightlike.* Since Δx^2 is a Lorentz invariant, a particular interval has the same quality (timelike, lightlike, or spacelike) when viewed in any Lorentz frame.

The Light Cone. Consider a spacetime diagram for particle b in which the x^2 and x^3 axes are suppressed. By convention, b is at $x^1 = 0$ when $x^0 = ct = 0$; it passes through the origin of this spacetime diagram. The path of b in this diagram is called its *world line.* Since the speed of b is always less than c, the slope of its world line is always greater than 45°.

Consider now an event a that occurs at some point in spacetime. As viewed in a spacetime diagram in the lab frame, relative to the present time $t = 0$ of particle b, event a may occur in the past ($t < 0$), present ($t = 0$), or future ($t > 0$). Since information cannot travel faster than the speed of light, only past events a that lie within the bottom *light cone* can already have affected b ("passive past" of b); conversely, present actions by b can affect only those future events a which lie within the top light cone ("active future" of b). More generally, in any Lorentz frame, if event a at $(ct_a, \vec{x}_a)$ and event b at $(ct_b, \vec{x}_b)$ are causally connected, the interval between them must be timelike or lightlike.

Time Dilation. If we are in the rest (primed) frame of a moving particle, $\Delta\vec{x}'$ is always 0 and $\Delta x'^2$ reduces to $c^2\Delta t'^2$. Since an observer in any inertial frame can use a Lorentz transformation to calculate the time interval that would be observed in the particle's rest frame, all such observers will agree on that particular time interval, which is called the *proper time* interval $\Delta\tau$. Then the proper time τ is a Lorentz invariant. However, because of this fact, it is not even necessary for an observer to perform a Lorentz transformation in order to calculate $\Delta\tau$. It is enough to calculate Δx^2 in the observer's *own* (lab) frame, obtaining

$$\begin{aligned} c^2\Delta\tau^2 &= (x'^0_a - x'^0_b)^2 \\ &= \Delta x^2 \\ &= c^2(t_a - t_b)^2 - \\ &- (x^1_a - x^1_b)^2 - (x^2_a - x^2_b)^2 - (x^3_a - x^3_b)^2 \ . \end{aligned}$$

From this equation it is clear that the time difference in the observer's frame is always $\geq$ the proper time. This is called "time dilation".

Exercise 3. Use the Lorentz transformation to show that in the lab frame the observed time difference Δt between two events that occur at the same 3-vector $\vec{x}'$ in a moving frame is a factor γ_0 *greater* than the proper time difference $\Delta\tau = \Delta t'$.

Space Contraction. Likewise show that the spatial separation Δx^1 observed in the lab between two events that occur at the same time t in the lab (*i.e.* measurement at the same laboratory time of the position of both ends of a moving rod) is a factor γ_0 *smaller* than the spatial separation $\Delta x'^1$ between the same two points in the moving system.

Rapidity. The rapidity η is defined by $\tanh\eta = \beta$. In particular, when β is the relative velocity β_0 between two Lorentz frames, η becomes η_0 and the Lorentz transformation is algebraically equivalent to

$$\begin{aligned} x^0 &= x'^0 \cosh\eta_0 + x'^1 \sinh\eta_0 \\ x^1 &= x'^0 \sinh\eta_0 + x'^1 \cosh\eta_0 \ . \end{aligned}$$

This looks like a rotation, except that hyperbolic rather than circular functions of the transformation parameter ("angle") are used. Intuitively you can imagine why. An ordinary rotation must preserve the length of an ordinary vector. This is achieved using the sin and cos functions, taking advantage of the identity $\sin^2\theta + \cos^2\theta = 1$. Lorentz transformations instead must preserve the length of 4-vectors, which involves the difference of the squares of the timelike and spacelike elements. Then the hyperbolic sin and cos are needed, since $\cosh^2\eta - \sinh^2\eta = 1$. So the rapidity plays the same role for Lorentz transformations as the angle plays for ordinary rotations. And Lorentz transformations are just rotations in spacetime.

Addition of Velocities. Write a standard Lorentz transformation

$$x^0 = \gamma_0(x'^0 + \beta_0 x'^1)$$

$$x^1 = \gamma_0(x'^1 + \beta_0 x'^0).$$

Then take the differential of it: $dx^0 = ...$; $dx^1 = ...$ Divide the top by the bottom equation and identify

$$\begin{aligned} \frac{dx^1}{dx^0} &= \beta = c^{-1} \times \text{ speed in lab frame} \\ \frac{dx'^1}{dx'^0} &= \beta' = c^{-1} \times \text{ speed in moving frame} \ . \end{aligned}$$

Then you easily obtain the Einstein law for the addition of velocities:

$$\beta = \frac{\beta' + \beta_0}{1 + \beta'\beta_0}.$$

Note that β can never exceed unity.

Exercise 4. Complete the above derivation. Also prove that the Einstein law reduces to

$\eta = \eta' + \eta_0$. Thus the rapidities *add linearly*. (Likewise, two successive Lorentz transformations in the same direction characterized by rapidities η_1 and η_2 are equivalent to one transformation characterized by rapidity $\eta_1 + \eta_2$.)

II. Other 4-Vectors

4-Velocity. We saw that the proper time τ of a moving particle is a Lorentz scalar, and $x = (ct, \vec{x})$ is a 4-vector. Then $dx/d\tau$ is also a 4-vector: it is the 4-velocity u. Suppose that the particle is moving in the $\hat{1}$ direction with velocity βc. Then, using time dilation,

$$u^0 = \frac{dx^0}{d\tau} = \frac{\gamma c d\tau}{d\tau} = \gamma c$$
$$u^1 = \frac{dx^1}{d\tau} = \frac{dx^1}{dt/\gamma} = \gamma\beta c \; .$$

Generalizing to particle motion with velocity $\vec{\beta}c$ in an arbitrary direction,

$$u^0 = \gamma c,\; \vec{u} = \gamma\vec{\beta}c, \text{ and } u \cdot u = c^2.$$

4-Momentum. The 4-momentum p is just mu, where m is the rest mass of the particle. Then $p = (\gamma mc, \gamma\vec{\beta}mc)$. (Please discard forever any previously learned notation in which m is the "relativistic mass" varying with velocity.) The 4-momentum is likewise a 4-vector obeying the Lorentz transformation, and its space components reduce to the familiar vector momentum $\vec{p}$ when $\beta \ll 1$. A necessary condition is that two identical particles in an elastic collision conserve total p in any frame, if total p is conserved in one frame. It is easy to show that our choice satisfies this condition. (But to prove that the choice is unique, it is necessary *e.g.* to apply this condition to Taylor series expansions of as-yet-unspecified energy-momentum functions).

Often $p = (p^0, \vec{p})$ is written $p = (E/c, \vec{p})$. $E = \gamma mc^2$ is the total energy and $\vec{p} = \gamma\vec{\beta}mc$ is the relativistic momentum. Then the condition $u \cdot u = c^2$ implies

$$E^2 - p^2c^2 = m^2c^4 \; ,$$

which is the well-known key to solving relativistic kinematics problems. The total energy E is further divided into rest-mass energy mc^2 and kinetic energy T. Then

$$E = \gamma mc^2 = T + mc^2;\; T = (\gamma - 1)mc^2 \; .$$

This result for the kinetic energy T reduces to the usual $\frac{1}{2}mv^2$ in the limit $\beta \ll 1$. A stunning consequence of the extra term mc^2 in the total energy is the possible conversion of rest-mass energy to kinetic energy.

Relativistic Doppler Shift. Taking advantage of your background in 137B, I begin with the relativistic de Broglie relation

$$E = \hbar\omega;\; \vec{p} = \hbar\vec{k}$$

where ω and $\vec{k}$ are the angular frequency and the wave vector of the wave. This is expressed more compactly as

$$p = \hbar k \; ,$$

where $k = (\omega/c, \vec{k})$ is the *wave 4-vector*.

Consider a wave source $\mathcal{S}$ that is at rest at coordinates $(0, x'^2, 0)$ in a frame moving with velocity $\hat{1}\beta_0 c$ with respect to an observer, who is at rest at the origin of the lab frame. As seen by the observer, the source is at an angle θ above the horizon ($\theta = 0$ if directly approaching, $\theta = \pi$ if directly receding). Using the inverse Lorentz transformation

$$k' = \Lambda(-\beta_0) \cdot k$$
$$k'^0 = \gamma_0 k^0 - \gamma_0\beta_0 k^1 \text{ where } k^1 = |\vec{k}|\cos\theta \; .$$

For a wave in a homogeneous non-dispersive medium,

$$|\vec{k}| = \frac{\omega}{v_{\text{phase}}} = \frac{k^0}{\beta} \; ,$$

where βc is its velocity of propagation in the medium (considered to be at rest in the laboratory). Putting it together,

$$k^0 = \frac{k'^0}{\gamma_0[1 - (\beta_0/\beta)\cos\theta]}.$$

Since $k^0 = \omega/c$, ω is related to ω' by the same formula. We obtain the familiar red shift ($\omega < \omega'$) for a receding source $\cos\theta < 0$, or a blue shift if the source is approaching almost directly ($\cos\theta \to 1$). The new result, not familar from a nonrelativistic treatment, is the extra red shift due to time dilation. For example, a red shift $\omega = \omega'/\gamma_0$ is present even when the source is *transverse* ($\theta = \pi/2$).

A similar analysis yields the relativistic relations for *aberration* of a wave

$$\tan\theta' = \frac{\sin\theta}{\gamma_0(\cos\theta - \beta_0\beta)} ,$$

and of a particle

$$\tan\theta' = \frac{\sin\theta}{\gamma_0(\cos\theta - \beta_0/\beta)} ,$$

where, for the particle, βc is its laboratory velocity.

Exercise 5. Derive the two relativistic relations for aberration.

III. Covariance of Electrodynamics

Contravariant Derivative. We need a 4-dimensional version of ∇ that transforms like a 4-vector. This is the *contravariant derivative*

$$\partial = (\frac{\partial}{c\partial t}, -\nabla).$$

The minus sign indeed is required in order to satisfy the property $\partial = \Lambda \cdot \partial'$; the reason has to do with the fact that x^μ appears in the denominator of ∂ rather than in the numerator. (Then the usual rule for taking dot products of 4-vectors causes $\partial \cdot x = 4$, not -2.)

4-Current Density. The usual equation for conservation of electric charge

$$\frac{\partial\rho}{\partial t} + \nabla \cdot \vec{J} = 0$$

can be rewritten

$$\partial \cdot J = 0 \text{ where } J = (c\rho, \vec{J}) .$$

Since charge is conserved in any reference frame, the 4-current density J must transform like a 4-vector.

4-Potential. The *Lorentz gauge condition*

$$\frac{\partial V}{c^2\partial t} + \nabla \cdot \vec{A} = 0$$

can be rewritten

$$\partial \cdot A = 0 \text{ where } A = (V/c, \vec{A}) .$$

Here V and $\vec{A}$ are the usual scalar and vector potentials. Again, for this gauge condition to be satisfied in any Lorentz frame, the 4-potential A must transform like a 4-vector. Of course, V and $\vec{A}$ may be chosen in a different gauge, for example Coulomb gauge. Then the 4-potential A would still exist, but it would bear a more complicated relationship to V and $\vec{A}$.

With the definitions of the contravariant derivative, the 4-current density, and the 4-potential, the wave equations for V and $\vec{A}$ in Lorentz gauge

$$(\nabla^2 - \frac{\partial^2}{c^2\partial t^2})V = -\frac{\rho}{\epsilon_0}$$
$$(\nabla^2 - \frac{\partial^2}{c^2\partial t^2})\vec{A} = -\mu_0\vec{J}$$

may be written even more beautifully as the single equation

$$(\partial \cdot \partial)A = \mu_0 J$$

where the operator

$$\partial \cdot \partial = (\frac{\partial^2}{c^2\partial t^2}, -\nabla^2)$$

is the negative of the *D'Alembertian.*

Recall that the magnetic and electric fields $\vec{B}$ and $\vec{E}$ are given by

$$\vec{B} = \nabla \times \vec{A}\, ;\ \vec{E} = -\nabla V - \frac{\partial\vec{A}}{\partial t} .$$

Now that you know how $A = (V/c, \vec{A})$ transforms, in principle you know how $\vec{B}$ and $\vec{E}$ transform. However, it is easier to proceed by constructing one additional quantity, the *field strength tensor*.

Field Strength Tensor. The field strength tensor $\mathcal{F}$ is defined, for each of its 16 elements, by the 16 equations

$$\mathcal{F}^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \ .$$

Since $\mathcal{F}$ is obviously antisymmetric in μ and ν, only 6 of its elements can be independent. The advantages of defining it are: (*i*) within constants, its elements are the 6 components of the magnetic and electric fields; (*ii*) because of its structure, the transformation properties of $\mathcal{F}$ are readily deduced.

Exercise 6. Using the definitions of ∂ and A, show by explicit calculation, element by element, that the field strength tensor is

$$\mathcal{F} = \begin{pmatrix} 0 & -E^1/c & -E^2/c & -E^3/c \\ E^1/c & 0 & -B^3 & B^2 \\ E^2/c & B^3 & 0 & -B^1 \\ E^3/c & -B^2 & B^1 & 0 \end{pmatrix} .$$

Exercise 7. Define the *covariant derivative* $\tilde{\partial}$ as

$$\tilde{\partial} = (\frac{\partial}{c\partial t}, +\nabla).$$

Prove that

$$\tilde{\partial} \cdot \mathcal{F} = \mu_0 J$$

where, as usual, J is the 4-current density. (To take the product of $\tilde{\partial}$ with $\mathcal{F}$, write $\tilde{\partial}$ and J as row matrices and use ordinary matrix multiplication. The above relation is a Lorentz-invariant way to write the two Maxwell equations that contain sources.)

Transforming the Electromagnetic Field. You may have noticed that both of the indices μ and ν of $\mathcal{F}^{\mu\nu}$ are on the top, rather than μ on the top and ν on the bottom as on page 1. Relative to the prescription on page 1, this causes a slight change in the method for taking the product of $\mathcal{F}$ with a 4-vector (see Exercise 7), but it is not otherwise a concern, and we will not pursue this detail further here.

Because of its simple construction we can easily see how $\mathcal{F}$ transforms. Consider a tensor $\mathcal{T}$ that is formed from two 4-vectors A and B according to $\mathcal{T}^{\mu\nu} = A^\mu B^\nu$. We want to find $\mathcal{T}'^{\mu\nu}$ such that it is equal to $A'^\mu B'^\nu$, where A' and B' are the 4-vectors transformed to a moving frame. Write

$$\begin{aligned} \mathcal{T}^{\mu\nu} &= (\Lambda \cdot A')(\Lambda \cdot B') \\ \mathcal{T}^{\mu\nu} &= (\Lambda^\mu_\rho A'^\rho)(\Lambda^\nu_\sigma B'^\sigma) \ . \end{aligned}$$

(In the second equation the matrix multiplication is shown explicitly. It is understandable when you realize that you are to sum over the repeated indices ρ and σ.) Rewriting a bit,

$$\begin{aligned} \mathcal{T}^{\mu\nu} &= \Lambda^\mu_\rho A'^\rho B'^\sigma \Lambda^\nu_\sigma \\ \mathcal{T}^{\mu\nu} &= \Lambda^\mu_\rho \mathcal{T}'^{\rho\sigma} \Lambda^\nu_\sigma \ . \end{aligned}$$

As Λ is symmetric, the last equation is a prescription for ordinary multiplication of three matrices,

$$\mathcal{T} = \Lambda \cdot \mathcal{T}' \cdot \Lambda.$$

Because it is the sum of two tensors formed in the same way, the field strength tensor $\mathcal{F}$ transforms using this same simple matrix equation.

Exercise 8. Substitute the actual elements of the Lorentz transformation matrix Λ and the field strength tensor $\mathcal{F}$ in this simple matrix equation. Prove that

$$\begin{aligned} E^1 &= E'^1 \\ B^1 &= B'^1 \\ \vec{E}_\perp &= \gamma_0(\vec{E}'_\perp - \vec{\beta}_0 \times c\vec{B}'_\perp) \\ c\vec{B}_\perp &= \gamma_0(c\vec{B}'_\perp + \vec{\beta}_0 \times \vec{E}'_\perp) \ . \end{aligned}$$

(These are the standard equations for the relativistic transformation of electric and magnetic fields. The subscript "$\perp$" refers to the (vector) component perpendicular to $\vec{\beta}_0$, which here is in the $\hat{1}$ direction. Of course, if $\vec{\beta}_0$ is in an arbitrary direction, the first two equations can be generalized by replacing the superscript "^{1}" by the subscript "$_\parallel$", to indicate the direction parallel to $\vec{\beta}_0$.)
